TITLE:

確率制御問題と最適停止時刻問題の関連について(計画数学とその周辺)

AUTHOR(S):

安田, 正実

CITATION:
安田, 正実. 確率制御問題と最適停止時刻問題の関連について(計画数学とその周辺). 数理解析研究所講究録 1987, 611: 90-98

ISSUE DATE:
1987-02

URL:
http://hdl.handle.net/2433/99761

RIGHT:

京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

数理解析研究所講究録
611巻 1987年 90-98

確率制御問題と最適停止時刻問題の関連について

千葉大学教養部　安田正実 (Masami Yasuda)

確率制御問題の中で，ここでは，確率特異制御 singular control, あるいは impulse control とよばれる問題を意味する．この問題は既に1966年の Bather/Chernoff の論文に始まり，以来多くの研究がなされている．最近，Karatzas/Shreve (1984,1985) などにより，最適停止時刻問題との関係が発表された．そのうちの一つの結果によると，特異制御の最適値を初期状態の関数として微分すれば，その導関数は最適停止時刻問題の最適値に一致する，という．あたかも数理計画問題における Geoffrion の微分安定性に対応する性質があるかのように筆者には思えた．しかし，これは，単に両者の問題に対するそれぞれの最適政策が同じ単調性構造，すなわち，状態の関数とする政策が単調である，という対応関係に過ぎないという結論である．さらにもう一つここで注意したいことは，停止時刻問題において考えるものと比べると特別なモデルしか扱っているに過ぎないから，ごく普通の場合に対応する拡張を考えた．しかし残念ながら，もとの関係はこの場合には成立せず，この２つの値が一致したのは，特別な線形モデルに限られることがわかった．最後に，停止問題に付随して定まる特異制御の双対定理を考えるとする．

1．確率特異制御と停止時刻問題

確率特異制御問題とは次の形である： 拡散過程 $X_t, t \geq 0$,

$$dX_t = \mu dt + \sigma d\omega_t, \quad \omega_t \text{ は Brown Motion},$$

μ, σ は定数，が与えられ，制御としては２つの右連続，非減少で $\{X_t\}$-可測な確率過程 $R = \{R_t\}$, $L = \{L_t\}$ を考える．

(1) $$Z_t := X_t + R_t - L_t, \quad t \geq 0, \quad Z_0 := x$$

によって定まる制御過程 Z_t を制御システム・ダイナミックスとして，目的関数は

(2) $$J(x;R,L) := E_x[\int_0^\infty e^{-\beta t}\{h(Z_t)dt + r\, dR_t + l\, dL_t\}] \longrightarrow \min_{R,L}$$

とする．ここで $-\infty < r, l < \infty$ は定数，$h(\cdot)$ は滑らかな凸関数で，$|x| \longrightarrow \infty$ のとき $h(x) \longrightarrow \infty$ とする．$0 < \beta < \infty$ は割引率．

この制御問題の最適値を

(3) $$f(x) := \min_{R,L} J(x;R,L)$$

とおくと，最適方程式は

(4) $$\min\{A f(x) - \beta f(x) + h(x),\ f'(x) + r,\ l - f'(x)\} = 0$$

で与えられる(Harrison/Taksar)．ただし $A = \mu\, d/dx + \sigma^2/2\ d^2/dx^2$.

特異制御とよばれる由来は，最適制御 R,L は右連続であるが，ルベーグ測度０の点で増加をするからである． Impulse control とよばれるのもインパルスが同じ振舞いをするからである(Bensoussan/Lions)． Reflected Follower Problem とよばれる場合もある．２つの組 (R,L) を制御としたが，一つの制御のみしか考えない場合を，Reflected に対して， Monotone Follower Problem とよばれる (Benes/Shepp

/Witsenhaussen). 歴史的には Monotone case から Reflected case へと移行されてきた. Reflected のもう1つの由来は, Brown Motion における反射壁を考える Reflected Problem の不連続版に相当していることもある. 最適性の証明には, 変分不等式の問題と定式化して semi-martingale の変換公式（一般化された Ito 公式）を用いるものが多い. 詳しくは参考文献を参照されたい.

さて上で述べた確率制御に対して, 最適停止時刻問題との関連を導こう. 反射壁ブラウン運動による証明があるが(Karatzas/Shreve), ここでは最適方程式の変形により直接関連を導く. まず制御問題の最適方程式 (4) を変形する. そのために目的関数 (2) の係数において

(5) $$0 < r + 1$$

を仮定する. いま最適方程式 (4) で表れた式の一部を $F(x) := Af(x) - \beta f(x) + h(x)$ とおくと, $F(x) \geqq 0$, $F(x) \longrightarrow \infty$ $(|x| \longrightarrow \infty)$, $\{x;\ F(x)=0\} = \{x;\ f'(x)+r>0,\ 1-f'(x)>0\} = \{x;\ F'(x)=0,\ f'(x)+r>0,\ 1-f'(x)>0\}$ である. つまり, $f'(x) + r > 0$, $1 - f'(x) > 0$ のときは

$$Af'(x) - \beta f'(x) + h'(x) = 0.$$

また $Af'(x) - \beta f'(x) + h'(x) \neq 0$ のときは, $F(x) \neq 0$ より,

$$f'(x) + r = 0 \quad \text{あるいは} \quad 1 - f'(x) = 0$$

で, (5) の仮定によって $r \neq -1$ であるから, これが同時に等号が成り立つことは起らない. このことは後の零和ゲームでも少しコメントする.

以上によって

$$\left.\begin{array}{r} Af'(x) - \beta f'(x) + h'(x) \neq 0 \\ 1 - f'(x) > 0 \end{array}\right\} \text{ のときは } r + f'(x) = 0$$

$$\left.\begin{array}{r} Af'(x) - \beta f'(x) + h'(x) \neq 0 \\ r + f'(x) > 0 \end{array}\right\} \text{ のときは } 1 - f'(x) = 0$$

$$\left.\begin{array}{l} 1 - f'(x) > 0 \\ r + f'(x) > 0 \end{array}\right\} \text{ のとき } Af'(x) - \beta f'(x) + h'(x) = 0$$

が成立する．　これらの式をまとめるために，２×２行列 $\{a_{ij}\}$ に対する零和（行列）ゲームの記号を導入しよう．零和ゲームにおいて純戦略が最適戦略ならば

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij}$$

であるから，この等しい値を $\mathrm{val}[a_{ij}]$ と表そう．実際，

$$a_{12} > a_{22} > a_{21}$$

ならば，任意の a_{11} に対して，零和ゲームは純戦略が最適となる．したがって仮定 (5) がこのことの成り立つための十分条件であって，これに基づく結果が上で示された式である．すなわち

(6) $$v(x) := f'(x)$$

とおけば，$1 - v(x) > 0 > -r - v(x)$ であるから，純戦略が最適であって

(7) $$\mathrm{val}\begin{bmatrix} Av(x) - \beta v(x) + h'(x) & 1 - v(x) \\ -r - v(x) & * \end{bmatrix} = 0$$

なる形にまとめて書き表される．ここで要素(2,2)は空 ＊ （起こらない事象）とするが，$-r - v(x)$ と $1 - v(x)$ の間の値ならば任意でよい．零和ゲームにおける２人のプレイヤーは，それぞれ stop, continue と

いう戦略を自由にとることが許されているが，　仮定(5) の下では，　同時に　stop という戦略を執らないことを要請している．　これが空になることの意味を表している．

このようにして我々は確率制御問題 (2) の最適値 (3) を微分した関数 (6) の最適方程式を，(7) の形に書くことができた．　この最適方程式は次で述べる最適停止時刻問題のゲーム変形の場合(Yasuda)　に他ならないものであることが分かる．

つぎに対応する停止問題を記述する．X_t を拡散過程として２つのマルコフ時刻 λ, μ に対して

(8.1)

$$\bar{v}(x) := \inf_{\mu} \sup_{\lambda} E_x \left[\int_0^{\min(\lambda,\mu)} e^{-\beta t} h'(X_t)dt - re^{-\beta\lambda} 1_{\{\lambda \leqq \mu\}} + le^{-\beta\mu} 1_{\{\lambda > \mu\}} \right]$$

(8.2)

$$\underline{v}(x) := \sup_{\lambda} \inf_{\mu} E_x \left[\int_0^{\min(\lambda,\mu)} e^{-\beta t} h'(X_t)dt - re^{-\beta\lambda} 1_{\{\lambda \leqq \mu\}} + le^{-\beta\mu} 1_{\{\lambda > \mu\}} \right]$$

を目的関数とする最適停止時刻問題のゲーム変形を考える．一般には

$$\bar{v}(x) \geq \underline{v}(x)$$

である．しかし，

<u>定理．</u>　仮定(5) があれば，

(9) $$\underline{v}(x) = \bar{v}(x)$$

が成り立ち，これを v(x) とすると先の最適方程式(7)を満たす．

この証明は省略するが，後で述べるもっと一般のモデルで成り立つ特別

な場合であって自明なことである．したがってこの定理より，最適停止時刻問題 (7) と 特異制御 (4) におけるそれぞれの最適値は関係式 (6) で結ばれることがわかった．

停止時刻問題 (8) のモデルは 停止した時に得る payoff が定数であり，確率変動していない特別な場合である．一般に取り扱うモデルの目的関数は

$$\int_0^{\min(\lambda,\mu)} e^{-\beta t} h'(X_t)dt - r(X_t)e^{-\beta\lambda} 1_{\{\lambda \leqq \mu\}} + l(X_t)e^{-\beta\mu} 1_{\{\lambda > \mu\}}$$

であり，$0 < r(x) + l(x)$ ならば (9) の結論が得られる． しかしこのことから先の問題を $r(x)$, $l(x)$ に拡張して次が成り立ちそうに思えるが，実はそうではない．

(4)に対応する式：

$$\min\{Af(x) - \beta f(x) + h(x),\ f'(x) + r(x),\ l(x) - f'(x)\} = 0$$

において $v(x) = f'(x)$ とすれば，(7) に対応する式は成り立つが，特異制御 (2) の最適方程式とはならない． Bang Bang type が optimal になるには線形でないと一般には不成立である． 式 (4) は自由境界問題として扱うことが多いが，このような拡張では， もっと困難な，未知関数の変数のなかに未知定数が入る問題になってしまう． 筆者には解法を知らないし，特異制御には未知問題として残されている．したがって特異制御を考える問題は残念ながら， せいぜい時刻に依存した線形な係数に限られている現状である．

２．特異制御の双対問題

双対問題として次を考える． 拡散過程 $X_t, t \geq 0$ は前と同じで，制御 $R = \{R_t\}$， $L = \{L_t\}$ も同じとするが， システム・ダイナミックスは

(10) $$W_t := X_t - R_t + L_t, \quad t \geq 0, \quad W_0 := x,$$

とし，目的関数は，

(11) $$K(x;R,L) := E_x \left[\int_0^\infty e^{-\beta t} \{h(W_t)dt - r\, dR_t - l\, dL_t\} \right] \longrightarrow \max_{R,L}$$

とする． 最適値

(12) $$g(x) := \max_{R,L} K(x;R,L)$$

は，前と全く同じ議論によって，次の最適方程式：

(13) $$\max \{ A g(x) - \beta g(x) + h(x),\ -g'(x) - r,\ g'(x) - l \} = 0$$

を満たすことがわかる． このとき次の定理が成り立つ．

<u>定理．</u>

(14) $$f'(x) = \bar{v}(x), \qquad g'(x) = \underline{v}(x).$$

ここで $\bar{v}(x)$ と $\underline{v}(x)$ は (8) の最適停止問題において定義されたもの．

したがって この定理より，一般には

$$f'(x) \geq g'(x)$$

であるが， $0 < r + l$ ならば

$$f'(x) = g'(x)$$

であることが明らか．

（証明） 最適停止時刻問題 (8.1)，(8.2) のそれぞれに対応する方程式

を求めれば，

$$\min\max\begin{bmatrix} A\bar{v}(x) - \beta\bar{v}(x) + h'(x) & 1 - \bar{v}(x) \\ -r - \bar{v}(x) & * \end{bmatrix} = 0,$$

$$\max\min\begin{bmatrix} A\underline{v}(x) - \beta\underline{v}(x) + h'(x) & 1 - \underline{v}(x) \\ -r - \underline{v}(x) & * \end{bmatrix} = 0$$

となり，これは上で定義した特異制御問題（3）と（12）とに対する最適方程式（4）と（13）に一致するからである．

行列の空要素 ＊ （起こらない事象）を書くことが気になれば，

$$\max\{\min(A\underline{v}(x) - \beta\underline{v}(x) + h'(x),\ 1 - \underline{v}(x)\},\ -r - \underline{v}(x)\} = 0$$

などと書く方がよいかも知れない．

参考文献

確率特異制御について

Bather,J, and Chernoff,H., "Sequential Decisions in the control of a Spaceship", Proc. 5th Berkley Symp. Math. Stat. Prob. 3 (1966) 181-207.

Bensoussan,A. and Lions,J.L., "Nouvelles Methodes en Control Impulsionnel", Appl. Math. Optim. 1(1975) 289-312.

Benes,V.E., Shepp,L.A. and Witsenhausen,H.S., "Some solvable stochastic control problems", Stochastics, 4(1980) 39-83.

Harrison,J.M., Sellke,T. and Taylor,A.J., "Impulse control of Brownian Motion", Math. Opr. Res. 8 (1983) 454-466.

Karatzas,I. and Shreve,S.E., "Connection between optimal stopping and stochastic control, I:Monotone follower problems", SIAM J. Control Optim. 22(1984) 856-877.

Karatzas,I. and Shreve,S.E., "Connection between optimal stopping and stochastic control, II:Reflected follower problems", SIAM J. Control Optim. 23(1985) 433-451.

最適停止時刻問題のゲーム変形について

Dynkin,E.B.,"The game variant of the optimal stopping problem", Dokl. Akad. Nauk USSR, 185(1969) 241-288.

Neveu,J.,"Maringales a temps disret", Masson, 1972.

Bismut,J.M.,"Sur un probleme de Dynkin", Z. Wahr. v. Gebiete, 39 (1977) 31-53.

Lepeltier,J.P. and Maingueneau,M.A.,"Le Jeu de Dynkin en Theorie Generale Sans L'Hypothese de Mokobodski", Stochastics, 13(1984) 25-44.

Yasuda,M.,"On a randomized strategy in Neveu's stopping problem", Stoch. Proc. Appl. 21(1985) 159-166.

数理計画法の双対性と微分安定性について

Geoffrion,A.M.,"Duality in Nonlinear programming:A simplified application-oriented development", SIAM Review, 13(1971) 1-37.